Introduction: Face Recognition using Eigenfaces

Abstract

This project is able to recognize a person's face by comparing facial structure to that of a known person. This is achieved by using forward facing photographs of individuals to render a two-dimensional representation of a human head. The system then projects the image onto a "face space" composed of a complete basis of "eigenfaces." Because of the similarity of face shape and features from person to person, face images fall within a relatively small region of the image space and as such can be reproduced with less than complete knowledge of the image space. When new images are fed into this system it can identify the person with a high rate of success with the robustness to identify correctly even in the presence of some image distortions.

Introduction

Do I Know You?

The human capacity to recognize particular individuals solely by observing the human face is quite remarkable. This capacity persists even through the passage of time, changes in appearance and partial occlusion. Because of this remarkable ability to generate near-perfect positive identifications, considerable attention has been paid to methods by which effective face recognition can be replicated on an electronic level. Certainly, if such a complicated process as the identification of a human individual based on a method as non-invasive as face recognition could be electronically achieved then fields such as bank and airport security could be vastly improved, identity theft could be further reduced and private sector security could be enhanced.

Many approaches to the overall face recognition problem (The Recognition Problem) have been devised over the years, but one of the most accurate and fastest ways to identify faces is to use what is called the "eigenface" technique. The eigenface technique uses a strong combination of linear algebra and statistical analysis to generate a set of basis faces--the eigenfaces--against which inputs are tested. This project seeks to take in a

large set of images of a group of known people and upon inputting an unknown face image, quickly and effectively determine whether or not it matches a known individual.

The following modules will provide a walk through exactly how this goal is achieved. Since this was not the first attempt at automated face recognition it is important to see what other approaches have been tried to appreciate the speed and accuracy of eigenfaces. This is not a simple and straightforward problem, so many different questions must be considered as one learns about this face recognition approach.

With a basic understanding achieved it is time for the real stuff, the implementation of the procedure. This has been broken down into smaller, more manageable steps. First the the set of basis eigenfaces must be derived from a set of initial images (Obtaining the Eigenface Basis). With this basis known individuals can be processed in order to pepare the system for detection by setting thresholds (Thresholds for Eigenface Recognition) and computing matrices of weights (Face Detection Using Eigenfaces). Finally, with such a system in place, tests of robustness can be performed in order to determine what quality of input images are necessary in order for successful identification to take place (Results of Eigenface Detection Tests). In this way, relevant conclusions (Conclusions for Eigenface Detection) can be drawn about the overall efficacy of the eigenface recognition method.

The Problem of Face Recognition

Face recognition is a very interesting quandry. Ideally a face detection system should be able to take a new face and return a name identifying that person. Mathematically, what possible approach would be robust and fairly computationally economical? If we have a database of people, every face has special features that define that person. Greg may have a wider forehead, while Jeff has a scar on his right eyebrow from a rugby match as a young tuck. One technique may be to go through every person in the database and characterize it by these small features. Another possible approach would be to take the face image as a whole identity.

Statistically, faces can also be very similar. Walking through a crowd without glasses, blurry vision can often result in misidentifying someone, thus yielding an awkward encounter. The statistical similarities between faces gives way to an identification approach that uses the full face. Using standard image sizes and the same initial conditions, a system can be built that looks at the statistical relationship of individual pixels. One person may have a greater distance between his or her eyes then another, so two regions of pixels will be correlated to one another differently for image sets of these two people.

From a signal processing perspective the face recognition problem essentially boils down to the identification of an individual based on an array of pixel intensities. Using only these input values and whatever information can be gleaned from other images of known individuals the face recognition problem seeks to assign a name to an unknown set of pixel intensities.

Characterizing the dependencies between pixel values becomes a statistical signal processing problem. The eigenface technique finds a way to create ghost-like faces that represent the majority of variance in an image database. Our system takes advantage of these similarities between faces to create a fairly accurate and computationally "cheap" face recognition system.

Face Recognition Background

The intuitive way to do face recognition is to look at the major features of the face and compare them to the same features on other faces. The first attempts to do this began in the 1960's with a semi-automated system. Marks were made on photographs to locate the major features; it used features such as eyes, ears, noses, and mouths. Then distances and ratios were computed from these marks to a common reference point and compared to reference data. In the early 1970's Goldstein, Harmon and Lesk created a system of 21 subjective markers such as hair color and lip thickness. This proved even harder to automate due to the subjective nature of many of the measurements still made completely by hand.

A more automated approach to recognition began with Fisher and Elschlagerb just a few years after the Goldstein paper. This approach measured the features above using templates of features of different pieces of the face and them mapped them all onto a global template. After continued research it was found that these features do not contain enough unique data to represent an adult face.

Another approach is the Connectionist approach, which seeks to classify the human face using a combination of both range of gestures and a set of identifying markers. This is usually implemented using 2-dimensional pattern recognition and neural net principles. Most of the time this approach requires a huge number of training faces to achieve decent accuracy; for that reason it has yet to be implemented on a large scale.

The first fully automated system to be developed utilized very general pattern recognition. It compared faces to a generic face model of expected features and created a series of patters for an image relative to this model. This approach is mainly statistical and relies on histograms and the grayscale value.

Kirby and Sirovich pioneered the eigenface approach in 1988 at Brown University. Since then, many people have built and expanded on the basic ideas described in their original paper. We received the idea for our approach from a paper by Turk and Pentland based on similar research conducted at MIT.
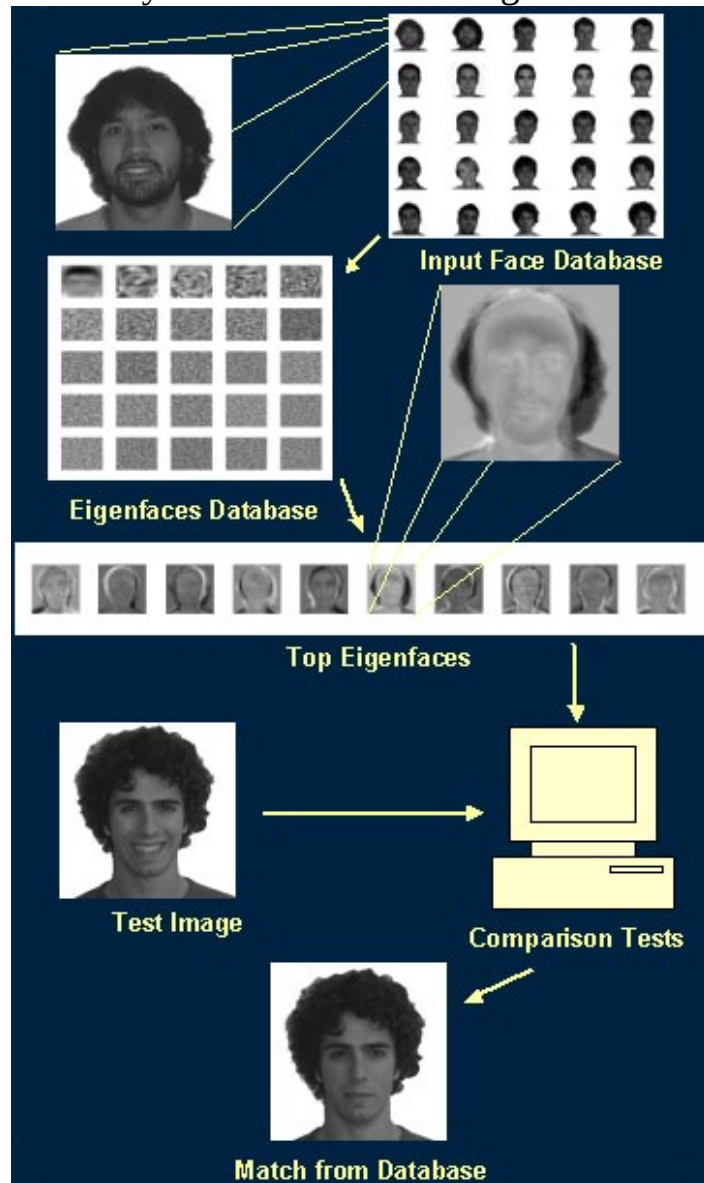
Obtaining the Eigenface Basis

Introduction to Eigenface System

The eigenface face recognition system can be divided into two main segments: creation of the eigenface basis and recognition, or detection, of a new face. The system follows the following general flow:
Summary of Overall Face Recognition Process

A robust detection system can yield correct matches when the person is

feeling happy or sad.

Deriving the Eigenface Basis

The eigenface technique is a powerful yet simple solution to the face recognition dilemma. In fact, it is really the most intuitive way to classify a face. As we have shown, old techniques focused on particular features of the face. The eigenface technique uses much more information by classifying faces based on general facial patterns. These patterns include, but are not limited to, the specific features of the face. By using more information, eigenface analysis is naturally more effective than feature-based face recognition.

Eigenfaces are fundamentally nothing more than basis vectors for real faces. This can be related directly to one of the most fundamental concepts in electrical engineering: Fourier analysis. Fourier analysis reveals that a sum of weighted sinusoids at differing frequencies can recompose a signal perfectly! In the same way, a sum of weighted eigenfaces can seamlessly reconstruct a specific person's face.

Determining what these eigenfaces are is the crux of this technique.

Before finding the eigenfaces, we first need to collect a set of face images. These face images become our database of known faces. We will later determine whether or not an unknown face matches any of these known faces. All face images must be the same size (in pixels), and for our purposes, they must be grayscale, with values ranging from 0 to 255. Each face image is converted into a vector Γ_n of length N (N=imagewidth*imageheight). The most useful face sets have multiple images per person. This sharply increases accuracy, due to the increased information available on each known individual. We will call our collection of faces "face space." This space is of dimension N.

Example Images from the Rice Database
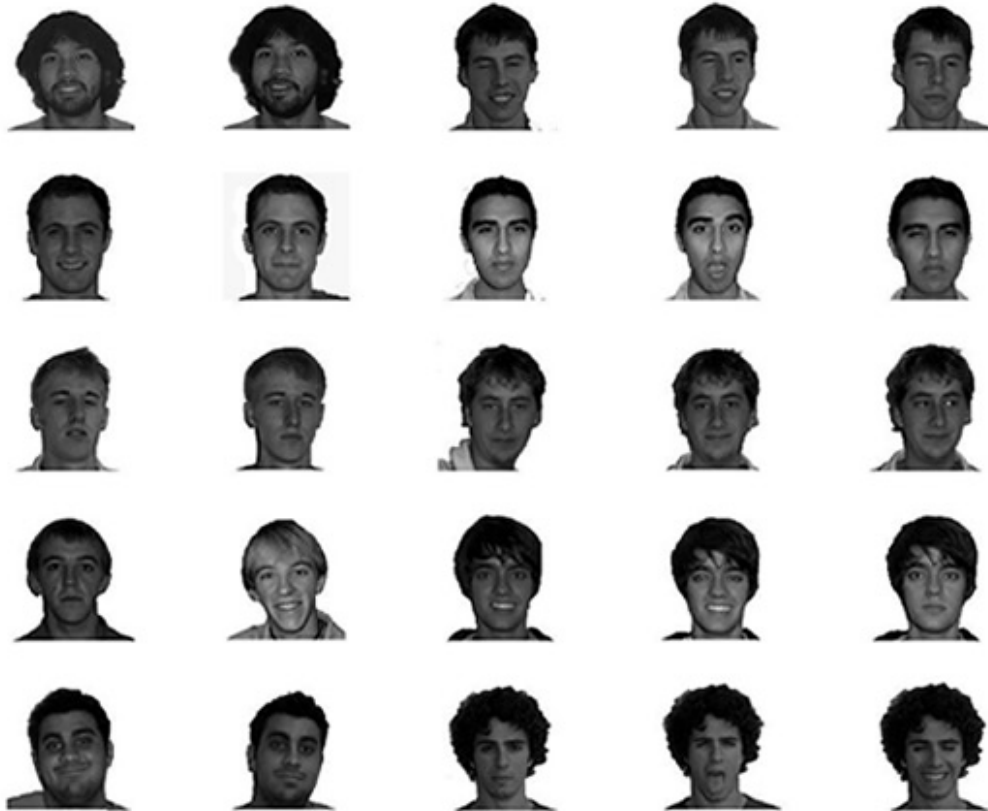
Next we need to calculate the average face in face space. Here M is the number of faces in our set:

Equation:

$$\Psi = \frac{1}{M} \sum_{n=1}^{M} \Gamma_n$$

Average Face from Rice Database

We then compute each face's difference from the average:
Equation:

$$\Phi_i = \Gamma_i - \Psi$$

We use these differences to compute a covariance matrix (C) for our dataset. The covariance between two sets of data reveals how much the sets correlate.
Equation:

$$C = \frac{1}{M} \sum_{n=1}^{M} \Phi_n \Phi_n^T = \frac{1}{M} \sum_{n=1}^{M} \begin{pmatrix} \text{var}(p_1) & \cdots & \text{cov}(p_1, p_N) \\ \vdots & \ddots & \vdots \\ \text{cov}(p_N, p_1) & \cdots & \text{var}(p_N) \end{pmatrix}_n = AA^T$$

Where $A = [\Phi_1 \Phi_2 ... \Phi_M]$ and p_i = pixel i of face n.

The eigenfaces that we are looking for are simply the eigenvectors of C. However, since C is of dimension N (the number of pixels in our images), solving for the eigenfaces gets ugly very quickly. Eigenface face recognition would not be possible if we had to do this. This is where the magic behind the eigenface system happens.

Simplifying the Initial Eigenface Basis

Based on a statistical technique known as Principal Component Analysis (PCA), we can reduce the number of eigenvectors for our covariance matrix from N (the number of pixels in our image) to M (the number of images in our dataset). This is huge! In general, PCA is used to describe a large dimensional space with a relative small set of vectors. It is a popular technique for finding patterns in data of high dimension, and is used commonly in both face recognition and image compression.* PCA is applicable to face recognition because face images usually are very similar to each other (relative to images of non-faces) and clearly share the same general pattern and structure.

PCA tells us that since we have only M images, we have only M non-trivial eigenvectors. We can solve for these eigenvectors by taking the eigenvectors of a new M x M matrix:

Equation:

$$L = A^T A$$

Because of the following math trick:

$$A^T A v_i = \mu_i v_i$$
$$A A^T A v_i = \mu_i A v_i$$

Where v_i is an eigenvector of L. From this simple proof we can see that $A v_i$ is an eigenvector of C.

The M eigenvectors of L are finally used to form the M eigenvectors u_l of C that form our eigenface basis:
**Equation:**

$$u_l = \sum_{k=1}^{M} v_{lk} \Phi_k$$

It turns out that only M-k eigenfaces are actually needed to produce a complete basis for the face space, where k is the number of unique individuals in the set of known faces.

In the end, one can get a decent reconstruction of the image using only a few eigenfaces (M'), where M' usually ranges anywhere from .1M to .2M. These correspond to the vectors with the highest eigenvalues and represent the most variance within face space.
Top Ten Eigenfaces from Rice Database
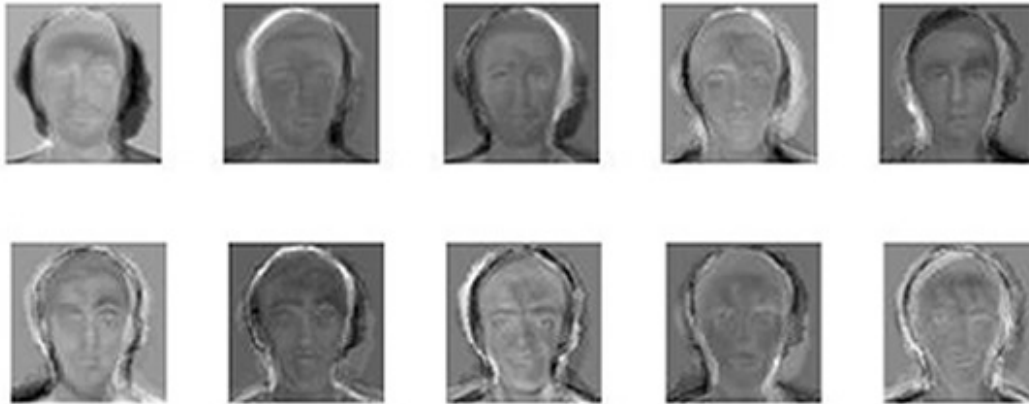
These eigenfaces provide a small yet powerful basis for face space. Using only a weighted sum of these eigenfaces, it is possible to reconstruct each face in the dataset. Yet the main application of eigenfaces, face recognition, takes this one step further.

*For more information on Principal Component Analysis, check out this easy to follow [tutorial](#).

Face Detection using Eigenfaces

Overview

Now that one has a collection of eigenface vectors, a question that may arise is, what next? Well, a sighted person can fairly easily recognize a face based on a rough reconstruction of an image using only a limited number of eigenfaces. However, reconstruction of non-face images is not so successful.
Poor Non-Face Reconstruction

I smell a rat, but certaintly not when I
reconstruct it with eigenfaces

Given that the initial objective is a face recognition system, eigenfaces happen to be a fairly easy, computationally economical, and successful method to determine if a given face is a known person, a new face, or not a face at all. A set of eigenface vectors can be thought of as linearly independent basis set for the face space. Each vector lives in its own dimension, and a set of M eigenfaces will yield an M dimensional space.

It should also be noted that the eigenfaces represent the principal components of the face set. These principal components are very useful in simplifying the recognition process of a set of data. To make it simpler, suppose we had a set of vectors that represented a person's weight and height. Projecting a given person onto these vectors would then yield that

person's corresponding weight and height components. Given a database of weight and height components, it would then be quite easy to find the closest matches between the tested person and the set of people in the database.

Equation:

$$w_p = Dot(Person, weight)$$
$$h_p = Dot(Person, height)$$

A similar process is used for face recognition with eigenfaces. First take all the mean subtracted images in the database and project them onto the face space. This is essentially the dot product of each face image with one of the eigenfaces. Combining vectors as matrices, one can get a weight matrix (M*N, N is total number of images in the database)

Equation:

$$\omega_k = \mu_k(\Gamma_{new} - \Psi)$$

Equation:

$$\Omega^T = [\omega_1 \omega_2 ... \omega_{M\prime}]$$

Equation:

$$WeightMatrix = \begin{bmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \vdots & \ddots & \vdots \\ \omega_{m\prime 1} & \cdots & \omega_{m\prime n} \end{bmatrix}$$

An incoming image can similarly be projected onto the face space. This will yield a vector in M dimensional space. M again is the number of used eigenfaces. Logically, faces of the same person will map fairly closely to one another in this face space. Recognition is simply a problem of finding the closest database image, or mathematically finding the minimum Euclidean distance between a test point and a database point.

Equation:

$$\varepsilon_k = \sqrt{||\Omega_{new} - \Omega_k||^2}$$

Due to overall similarities in face structure, face pixels follow an overall "face" distribution. A combination of this distribution and principal component analysis allows for a dimensional reduction, where only the first several eigenfaces represent the majority information in the system. The computational complexity becomes extremely reduced, making most computer programs happy. In our system, two techniques were used for image recognition.

Averaging Technique

Within a given database, all weight vectors of a like person are averaged together. This creates a "face class" where an even smaller weight matrix represents the general faces of the entire system. When a new image comes in, its weight vector is created by projecting it onto the face space. The face is then matched to the face class that minimizes the euclidean distance. A 'hit' is counted if the image matches correctly its own face class. A 'miss' occurs if the minimum distance matches to a face class of another person. For example, the ATT database has four hundred images total, composed of forty people with ten images each. The averaging technique thus yields a weight matrix with forty vectors (forty distinct face classes).

Removal Technique

This procedure varies only slightly from the averaging technique in one key way. The weight matrix represents the image projection vectors for images of the entire database. For empirical results, an image is removed from the system, and then projected onto the face space. The resulting weight vector is then compared to the weight vector of all images. The image is then matched to the face image that minimizes the euclidean distance. A 'hit' is counted if the tested image matches closest to another image of the same person. A 'miss' occurs when the image matches to any image of a different

person. The main difference from the average technique is the number of possible images that the test face can match to that will still result in a hit. For the ATT database, a weight matrix with four hundred vectors is used, but a new image could potentially 'hit' to ten distinct faces.

Thresholds for Eigenface Recognition

When a new image comes into the system, there are three special cases for recognition.

- Image is a known face in the database
- Image is a face, but of an unknown person
- Image is not a face at all. May be a coke can, a door, or an animal.

For a real system, where the pictures are of standard format like a driver's license photo, the first two cases are useful. In general, the case where one tries to identify a random picture, such a slice of pizza, with a set of faces images is pretty unrealistic. Nonetheless, one can still define these threshold values to characterize the images.

Looking back at the weight matrix of values using M eigenfaces, let's define the face space as an M-dimensional sphere encompassing all weight vectors in the entire database. A fairly approximate radius of this face space will then be half the diameter of this sphere, or mathematically, half the distance between the furthest points in the sphere.
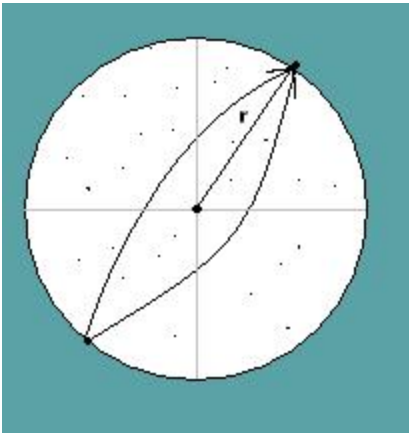
Equation:

To judge whether a new image falls within this radius, let's calculate the reconstruction error between the image and its reconstruction using M

eigenfaces. If the image projects fairly well onto the face space (image follows a face distribution), then the error will be small. However a non face image will almost always lie outside the radius of the face space.

Equation:

Equation:

Equation:

If the resulting reconstruction error is greater than the threshold, then the tested image probably is not a face image. Similar thresholds can be calculated for images of like faces. If a image passes the initial face test, it can be compared to the threshold values of faces in the database. A similar match process can be used as mentioned earlier. Also the removal or averaging technique can be applied for detection as previously described.

Results of Eigenface Detection Tests

Undistorted Input Results
For both the averaging technique and removal technique undistorted duplicates of the original images were processed for recognition in order to determine a best-case rate for recognition. For both techniques and for all three data sets, rates of recognition stabilized as the number of eigenfaces used in the recognition scheme increased.
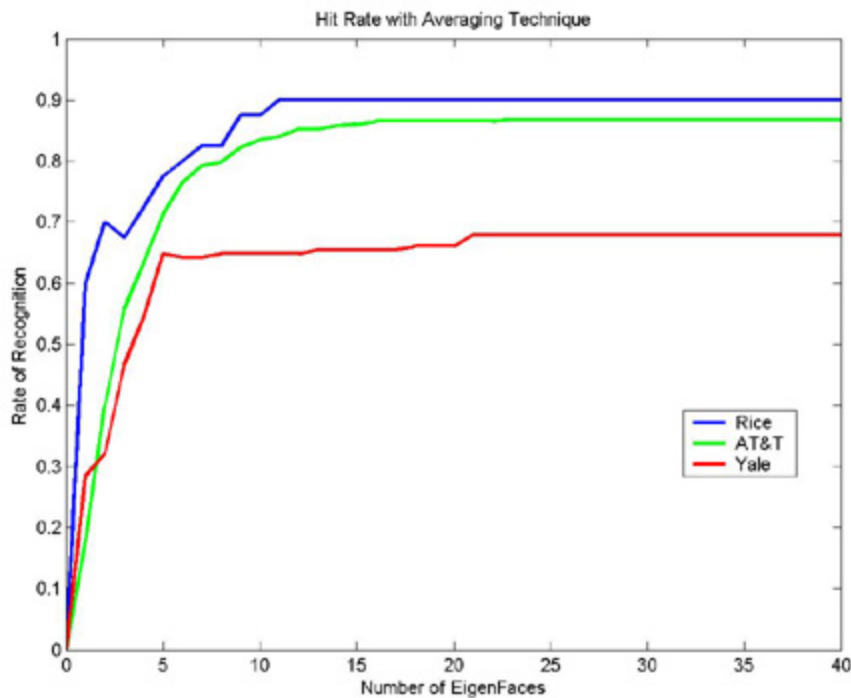
Rate of identification of the correct individual using undistorted inputs for the averaging technique.
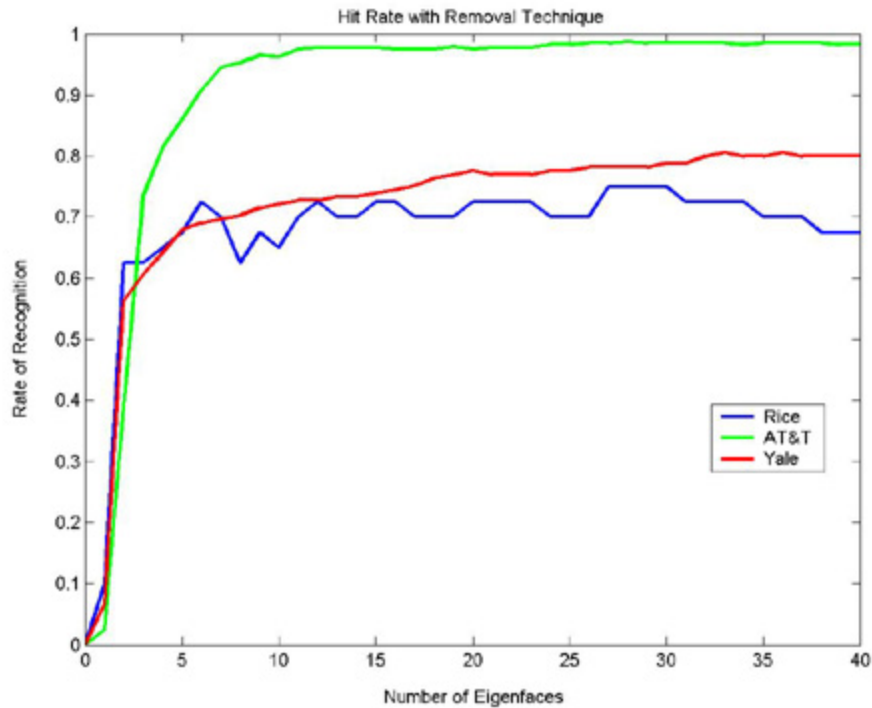
Rate of identification of the correct individual using
undistorted inputs for the removal technique.

For each image set stability was reached at the following hit rate and for the specified number of eigenfaces:

| Image Set	Stable Hit Rate	Number of Eigenfaces
Rice (Average)	90%	11
AT&T (Average)	86%	17
Yale (Average)	68%	21

Rice (Removal)	67%	14
AT&T (Removal)	96%	12
Yale (Removal)	75%	20

Table 1. Number of Eigenfaces for Hit Rate Stability for All Image Sets

For detection tests using a number of eigenfaces greater than that specified in Table 1 no significant improvement in detection success rate was achieved. In this way, undistorted tests suggest that implementations for both averaging and removal techniques do not achieve greater detection rates with numbers of eigenfaces greater than the minimum number needed for stability.

Occluded Input Results
For both averaging and removal techniques Rice image sets were tested for detection rates with horizontal and vertical occlusions centered on the vertical and horizontal axes respectively. Results show that hit rate stability, as before, is achieved as the number of eigenfaces used increases.
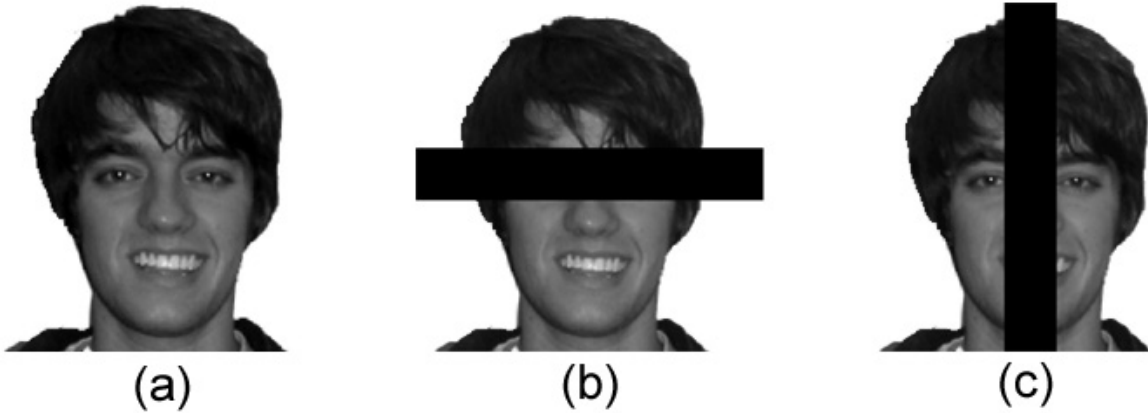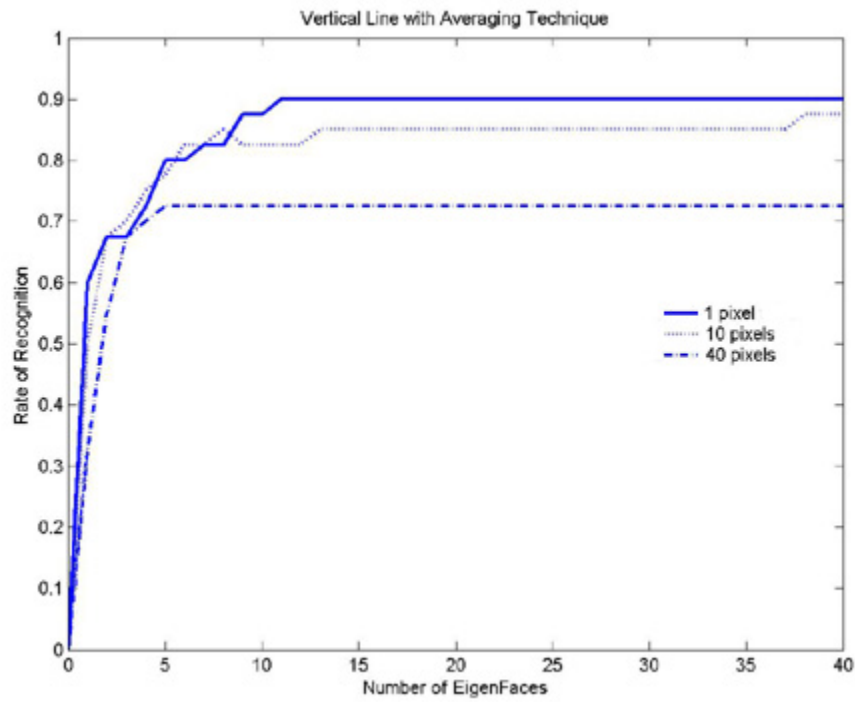

(a) the undistorted base image. (b) the image with a horizontal occlusion. (c) the image with a vertcal occlusion.
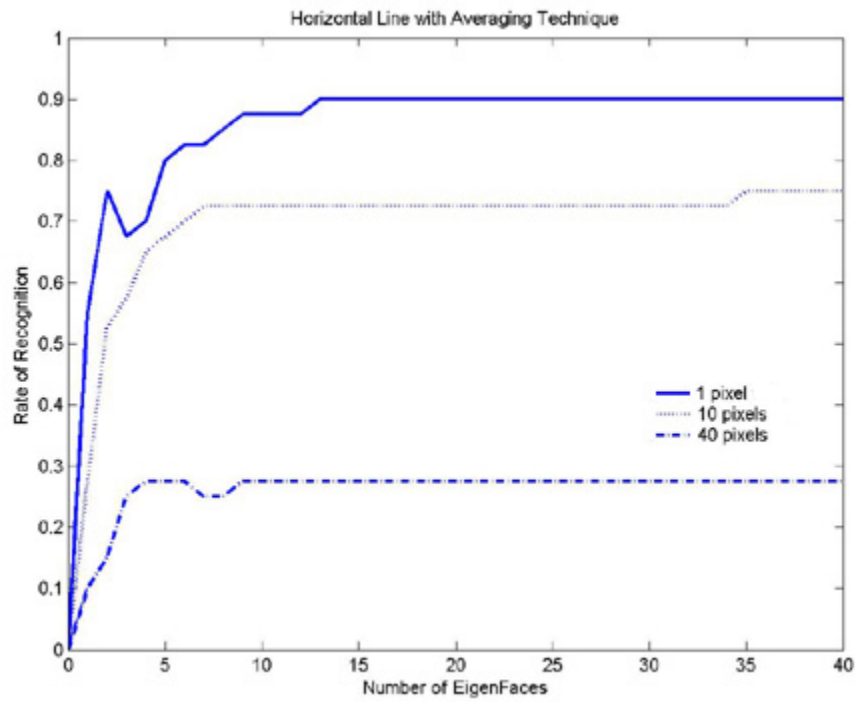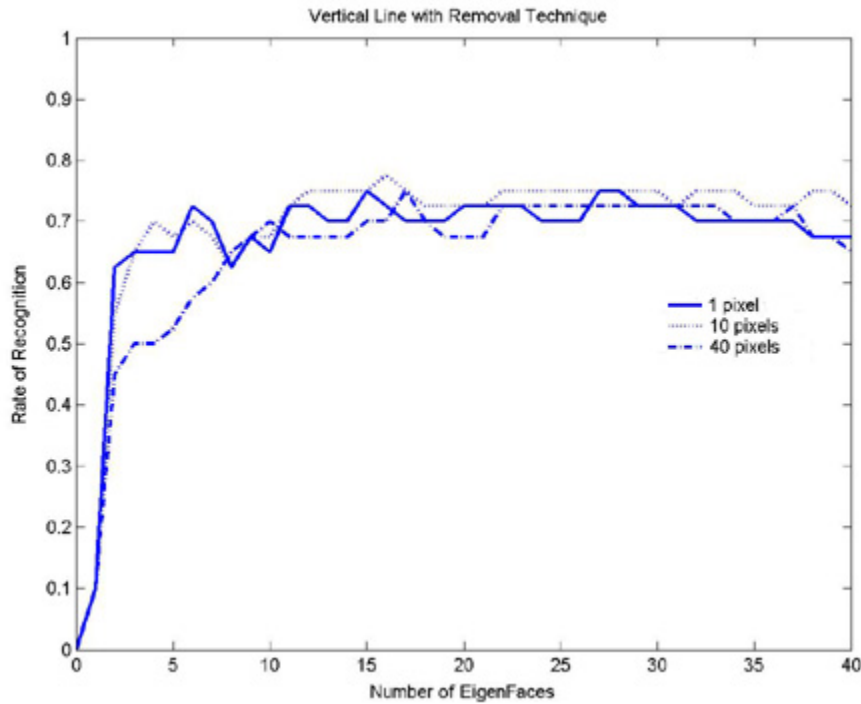
Rate of identification of the correct individual using horizontally obscured inputs with the averaging technique.
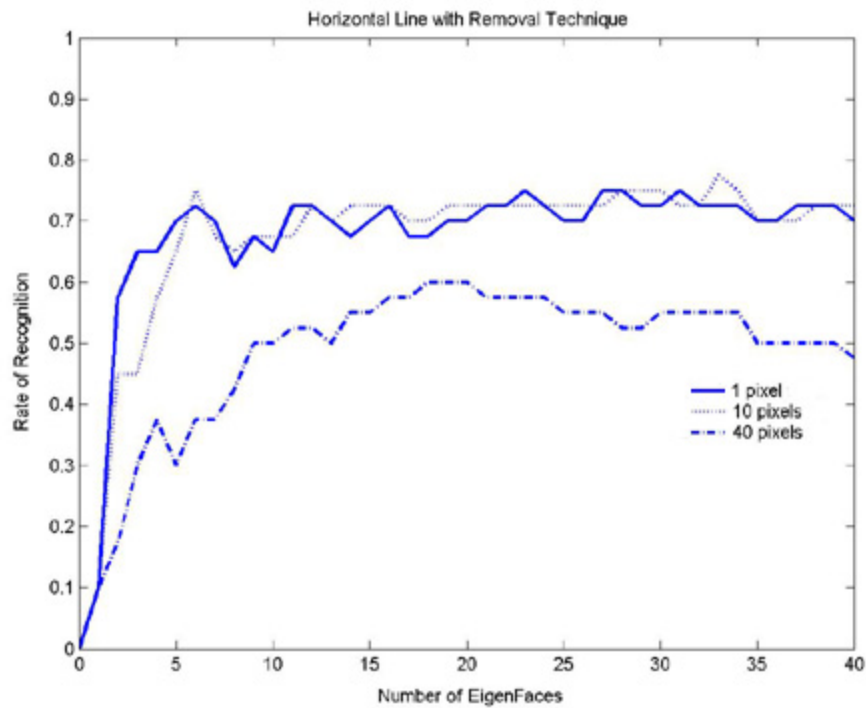
Rate of identification of the correct individual using horizontally obscured inputs with the removal technique.

Rate of identification of the correct individual using vertically obscured inputs with the averaging technique.

Rate of identification of the correct individual using vertically obscured inputs with the removal technique.

For each image set stability was reached at the following hit rate and for the specified number of eigenfaces:

Occlusion	Stable Hit Rate	Number of Eigenfaces
Horizontal 1 pixel	90%	13
Horizontal 10 pixels	72%	7
Horizontal 40 pixels	28%	4

Vertical 1 pixel	90%	11
Vertical 10 pixels	83%	7
Vertical 40 pixels	71%	5

Table 2. Number of Eigenfaces for Hit Rate Stability for Rice Image Set, Averaging Technique

Occlusion	Stable Hit Rate	Number of Eigenfaces
Horizontal 1 pixel	72%	11
Horizontal 10 pixels	72%	12
Horizontal 40 pixels	56%	14
Vertical 1 pixel	72%	11
Vertical 10 pixels	74%	12
Vertical 40 pixels	72%	15

Table 3. Number of Eigenfaces for Hit Rate Stability for Rice Image Set, Removal Technique

For detection tests using a number of eigenfaces greater than that specified in Tables 2 and 3 no significant improvement in detection success rate was achieved. In this way, occlusion tests suggest that implementations for both averaging and removal techniques do not achieve greater detection rates with numbers of eigenfaces greater than the minimum number needed for stability without occlusions.

Blurred Input Results

For both averaging and removal techniques, Rice image sets were tested for detection rates after being filtered with a two dimensional boxcar blur of various lengths. Results continue to indicate that the use of eigenfaces beyond the minimum necessary to achieve stability in the undistorted case is still unnecessary.
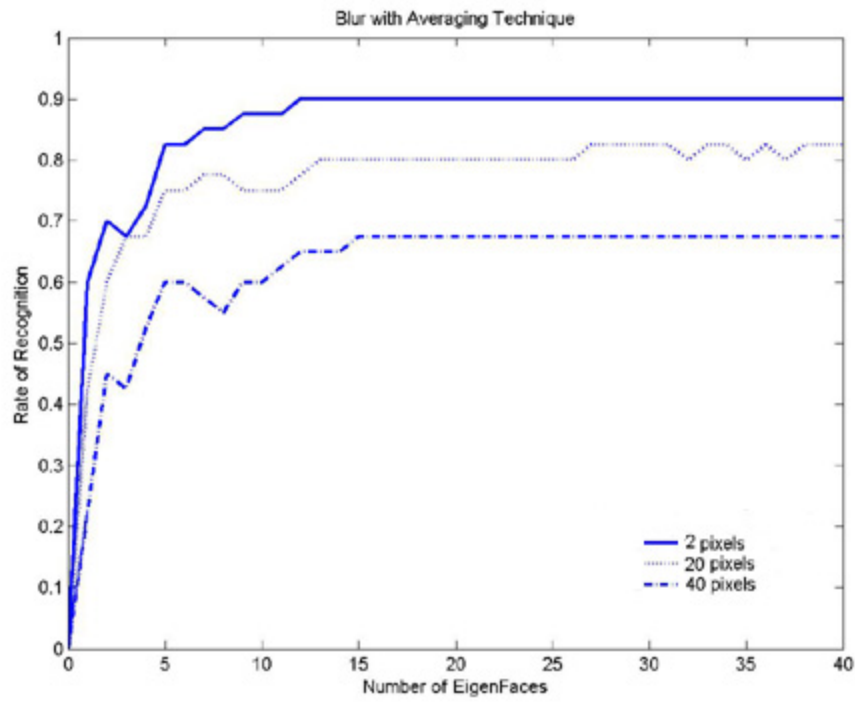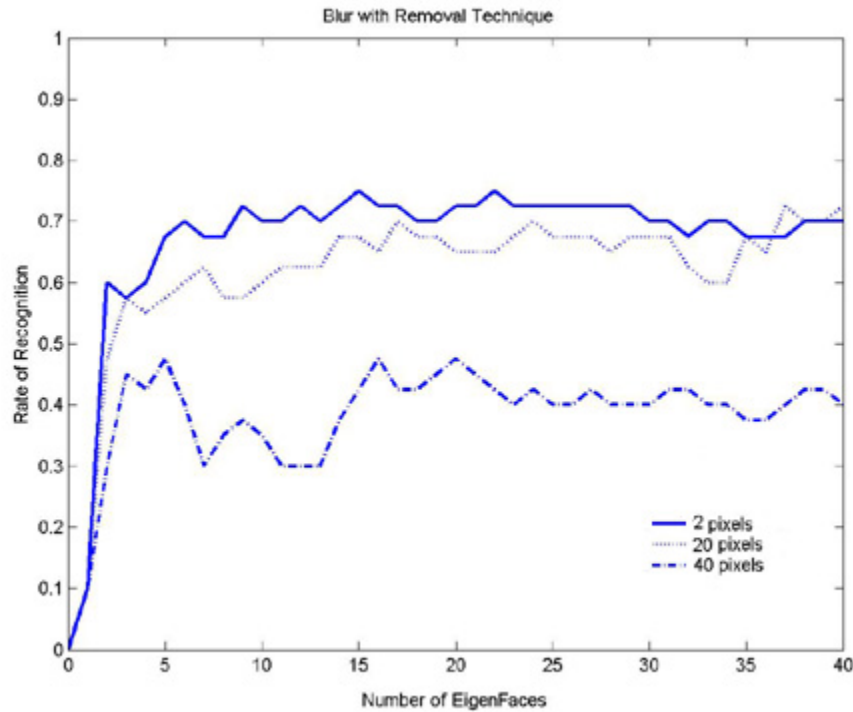

(a) undistorted base image. (b) same image with 20 pixel 2D boxcar blur.

Rate of identification of the correct individual for blurred input images with the averaging technique.

Rate of identification of the correct individual for blurred input images with the removal technique.

For each image set stability was reached at the following hit rate and for the specified number of eigenfaces:

Boxcar Length	Stable Hit Rate	Number of Eigenfaces
2 pixels	90%	12
20 pixels	80%	13
40 pixels	67%	15

Table 4. Number of Eigenfaces for Hit Rate Stability for Rice Image Set, Averaging Technique

Boxcar Length	Stable Hit Rate	Number of Eigenfaces
2 pixels	72%	9
20 pixels	66%	14
40 pixels	44%	15

Table 5. Number of Eigenfaces for Hit Rate Stability for Rice Image Set, Averaging Technique

For detection tests using a number of eigenfaces greater than that specified in Tables 4 and 5 no significant improvement in detection success rate was achieved. In this way, blurring tests suggest that implementations for both averaging and removal techniques do not achieve greater detection rates with numbers of eigenfaces greater than the number needed for stability without blurring.

Conclusions for Eigenface Detection

Analysis of the eigenface recognition technique using both averaging and removal methods gives evidence that the methods prove, at best, 90% accurate. In both cases, plateaus of recognition rates for a given number of eigenfaces are reached relatively quickly. This indicates that in any implementation of such a recognition system there does not exist a meaningful advantage to using more eigenfaces than first provide the desired level of accuracy. Furthermore, measurements of accuracy with various vertical and horizontal occlusions and two-dimensional boxcar blurs also demonstrate that excess eigenfaces provide no benefit in sub-optimal conditions.

In this way it becomes evident that if higher success rates are to be assured in most reasonable conditions then refinements to the eigenface concept must be made. Anecdotal experimentation with acquired image sets indicates that profile size, complexion, ambient lighting and facial angle play significant parts in the recognition of a particular image. Further research could be conducted into the viability of using eigenfaces and weightings taken for varying angles and lighting situations in order to allow for greater variability in both input images and detection opportunities. Clearly the eigenface offers much promise for the field of facial image recognition but not before some technical refinement.